

OPEN

# A Silicon Photonics Computational Lensless Active-Flat-Optics Imaging System

Alexander White\*, Parham Khial, Fariborz Salehi, Babak Hassibi &amp; Ali Hajimiri

The need for lightweight, miniature imaging systems is becoming increasingly prevalent in light of the development of wearable electronics, IoT devices, and drones. Computational imaging enables new types of imaging systems that replace standard optical components like lenses with cleverly designed computational processes. Traditionally, many of these types of systems use conventional complementary metal oxide semiconductor (CMOS) or charge coupled device (CCD) sensors for data collection. While this allows for rapid development of large-scale systems, the lack of system-sensor co-design limits the compactness and performance. Here we propose integrated photonics as a candidate platform for the implementation of such co-integrated systems. Using grating couplers and co-designed computational processing in lieu of a lens, we demonstrate the use of silicon photonics as a viable platform for computational imaging with a prototype lensless imaging device. The proof-of-concept device has 20 sensors and a 45-degree field of view, and its optics and sensors are contained within a  $2,000\ \mu\text{m} \times 200\ \mu\text{m} \times 20\ \mu\text{m}$  volume.

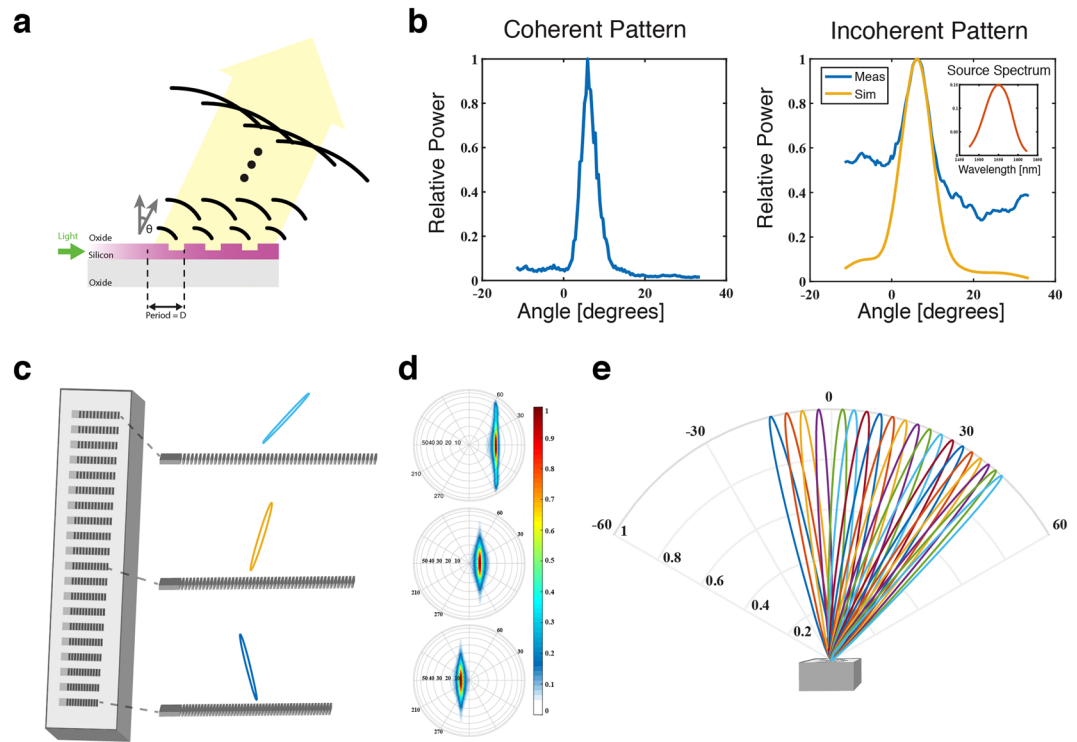
Lightweight miniature imaging systems are becoming essential to the development of wearable electronics, IoT devices, and drones<sup>1–3</sup>. Computation can be used to estimate information that is not captured by the imager, as in phase-retrieval imaging<sup>4</sup>, three dimensional imaging<sup>5</sup>, and encoded/coded aperture imaging<sup>6</sup>. Cleverly designed computational processes can also be used in imaging systems to replace standard optical components like lenses<sup>7</sup>. Traditionally, many of these types of systems use conventional CMOS or CCD sensors for data collection<sup>4,6,8</sup>. While the use of these prepackaged sensors allows for rapid development of large-scale systems, it fails to realize the full potential of a co-designed system-sensor in terms of versatility, performance, size, and mass<sup>3,9</sup>.

In this paper, we present an imaging system based on a custom-designed array of unique diffraction gratings (acting as optical receivers) co-designed with an adaptive reconstruction algorithm and implemented on an integrated photonics platform. This approach enables imaging within a very small volume, without the need for the depth typically required for most imaging systems. Integrated photonics allows for the capturing, manipulation, and sensing of light within an active layer only a few microns thick on a chip using dielectric waveguides manufactured with standard CMOS-type processing<sup>10</sup>. Furthermore, the built-in directionality of the custom integrated photonics microgratings can serve as a means of enhancing the computational image-recovery performance. These diffraction gratings can be engineered to have sensitivity in nearly arbitrary patterns and to admit a large span of wavelengths<sup>11,12</sup>, and as we show, can respond to incoherent light as well as coherent light. Using micrograting couplers co-designed with computational processing, we demonstrate that silicon photonics is a viable platform for computational imaging using a prototype lensless imaging device.

## Results

**Sensor design and implementation.** Images are representations of light field patterns where each point represents the intensity (and possibly other properties, such as color) of light impinging on the imager from a different bearing. When an image is displayed, these points replicate the original light, and the eye can process the image similarly to the original scene, using a lens to direct light from each angle to a different photoreceptor cell on the retina. Conventional cameras work in the same way, using a lens to redirect light impinging at different angles onto different pixels. However, this is not the only way to capture an image. As long as enough independent information is captured from a light field pattern, an image can be formed. For instance, this can be achieved, as in the retinal cells in the compound eye of an insect, with sensors/pixels individually sensitive to light impinging

Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA. \*email: [adwhite@stanford.edu](mailto:adwhite@stanford.edu)



**Figure 1.** (a) Operation of a silicon photonic grating coupler. Light guided in a silicon waveguide clad by silicon dioxide is scattered by regularly spaced etched gratings. These scattered waves interfere coherently in a direction  $\theta$  dependent on the period of gratings  $D$ . (b) The measured coherent and incoherent patterns of a grating coupler. The incoherent pattern is measured with the source spectrum shown. The simulated data is the theoretical pattern given the coherent pattern, the source pattern, and equation (5). (c) Visual representation of the fabricated device containing 20 uniquely spaced grating couplers. Three example couplers are shown with their simulated angular reception patterns. (d) Hemispherical projection of angular reception patterns of the corresponding grating couplers in c. (e) Simulated angular reception patterns of the full 20 grating couplers, in the direction of the gratings.

from different directions<sup>13,14</sup>, or even in a different domain, such as Fourier space<sup>15</sup>. While the data itself is not in the form of an image, various computation algorithms can be used to generate a recognizable image from the captured information<sup>7</sup>.

Here, we use an array of individually designed silicon-on-insulator (SOI) diffraction gratings to capture spatial intensity information from incoherent light with a wavelength centered at 1550 nm. By using an array of custom micrograting couplers, each with a different sensitivity profile vs. angle of incidence, we can capture enough independent information to linearly map our data into spatial pixels.

Micrograting couplers are versatile devices used to couple light from free space to guided modes in integrated photonics waveguides. Basic grating couplers consist of a periodic structure that diffracts incoming light such that it interferes coherently into the plane of the grating and a taper or other mode matching structure that guides the light into a waveguide mode, Fig. 1a. The periodicity and effective indexes of its internal components serve as design parameters that can be used to achieve the desired angular sensitivity profile of a coupler. For instance, in the case of uniformly spaced grating elements, the angle of maximum sensitivity is given by:

$$\theta_{max} = \sin^{-1} \left( n_{eff} - \frac{k\lambda_0}{d} \right) \quad (1)$$

where  $n_{eff}$  is the effective index of the grating structure,  $\lambda_0$  is the free space wavelength,  $d$  is the distance between gratings, and  $k$  is any integer.

The one-dimensional radiation pattern of a uniform coupler in the direction parallel to its main axis can be approximated by assuming an exponential optical loss along the coupler in conjunction with the propagation phase shift:

$$h(\theta) \propto \sum_k (1 - \alpha)^{k-1} e^{i \left( \frac{d}{\lambda_{eff}} - \frac{d}{\lambda_0} \sin(\theta) \right)} \quad (2)$$

where  $\alpha$  is the loss per grating tooth and  $\lambda_{eff}$  is the effective wavelength inside the grating structure. Given such a pattern  $h(\theta)$ , the power received at time,  $t$ , due to an arbitrary farfield illumination  $I(\theta, t)$  can be given as

$$P(t) = \left| \int_{\theta} h(\theta) I(\theta, t) d\theta \right|^2 \quad (3)$$

Under spatially incoherent illumination,  $I(\theta)$  and  $I^*(\theta')$  are uncorrelated except when  $\theta = \theta'$ . If the illumination is also temporally incoherent, we can represent the total power as

$$P = \int_{\theta} \int_{\omega} \left| h(\theta, \omega) \right|^2 \left| I(\theta, \omega) \right|^2 d\omega d\theta \quad (4)$$

with a single illumination source, we can separate the angular and frequency dependence of  $I$ , and define a modified sensitivity pattern  $h'(\theta)$ ,

$$\left| h'(\theta) \right|^2 \equiv \int_{\omega} \left| h(\theta, \omega) \right|^2 S(\omega) d\omega \quad (5)$$

where  $S(\omega)$  is the frequency spectrum of the scene illumination. Under narrow-band illumination, the pattern of a uniform coupler can be approximated as the convolution of the coherent pattern and the optical power spectrum, Fig. 1b.

For non-uniform couplers or extremely broadband illumination, the sensitivity pattern can be calculated by explicit simulation and weighted averaging of the patterns and the wavelength component of the optical power spectrum. As incoherent illumination leads to power addition and not electric field addition (as it would under coherent illumination), the power received by an arbitrary far-field distribution is given by the inner product of the sensitivity pattern and the far-field pattern. We can discretize this inner product to have the following approximation,

$$P \approx \sum_{\theta} \left| h'(\theta) \right|^2 \left| I(\theta) \right|^2 \Delta\theta, \quad (6)$$

where the number of discretization points determines the accuracy of our estimation. Consequently, we can rewrite (6) as the following,

$$P = \mathbf{h}^T \mathbf{x}, \quad (7)$$

where  $\mathbf{h}$  is an  $n$  dimensional positive pattern vector ( $\mathbf{h} \in \mathbb{R}_+^n$ ), and  $\mathbf{x}$  is the intensity vector, where  $n$  is given by the number of the distinct values of  $\theta$  we have used in the approximation (6). If we combine the results from multiple sensors, we can get the following linear system of equations,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \quad (8)$$

where the  $m$ -dimensional positive vector  $\mathbf{y}$  is the measured power,  $m$  is the number of sensors, the  $m$ -dimensional

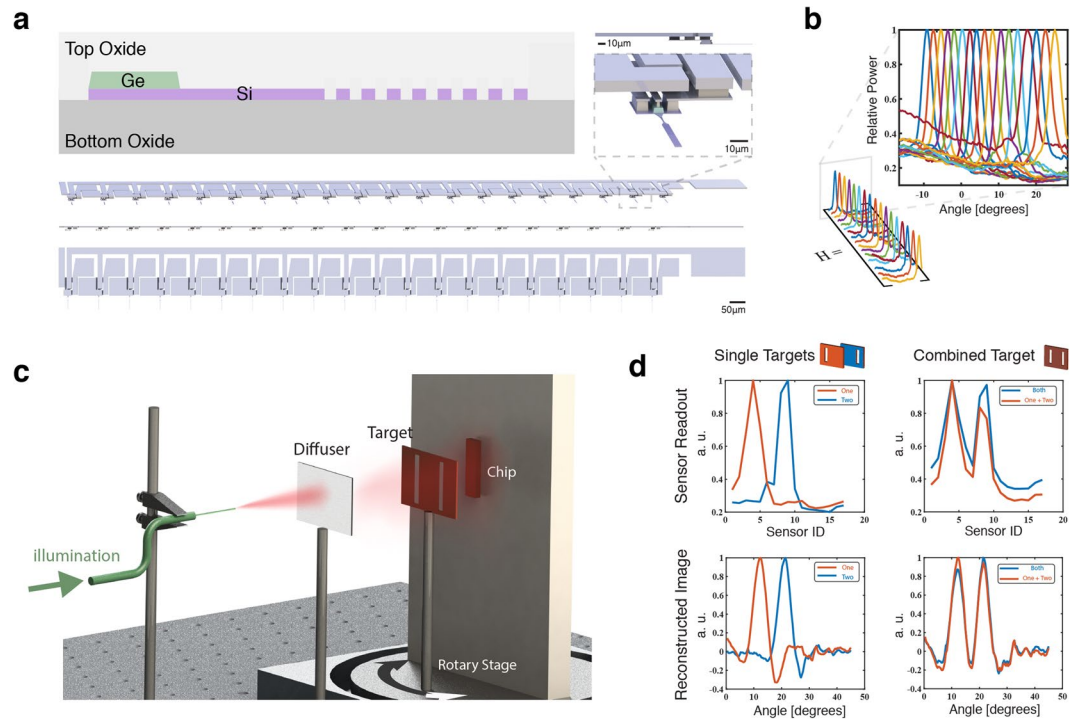
vector  $\mathbf{z}$  is the unknown noise vector, and  $\mathbf{H} = \begin{bmatrix} \mathbf{h}_1^T \\ \mathbf{h}_2^T \\ \vdots \\ \mathbf{h}_m^T \end{bmatrix}$  is the matrix whose rows correspond to the pattern of each

sensor. As we construct  $\mathbf{H}$  through the design of our grating couplers, we can reconstruct the image  $\mathbf{x}$  by inverting the linear system.

To form an imager, we can modulate the grating pitch across a number of couplers, effectively generating a set which is sensitive to different illumination angles. In this way, the sensors mimic the effect of a lens on traditional pixels. As each grating coupler is designed to be most sensitive to a different angle, the approximated transfer matrix  $\mathbf{H}$  (between discrete angles and sensor power) can be inverted effectively. Heuristically, overlapping sensor patterns allow for the data capture and localization of the full imaging area, and single angle reception allows for more independent measurements, improving the resolution or minimum resolvable angle. Thus it is favorable to strike a balance between overlapping and narrow sensor patterns.

As a proof of concept, we designed and fabricated a device with 20 of these sensors on a silicon photonics SOI platform, Fig. 1c. Each sensor consists of a uniform grating coupler and a germanium photodiode, sensitive to light with a wavelength around 1,550 nm. The grating couplers guide impinging light into 500 nm by 220 nm transverse-electric single mode silicon waveguides, which are then fed to the germanium photodiodes. The photodiodes convert the optical power received to an electric current, which can be amplified and measured. These sensors have a far-field coherent full-width half-maximum sensitivity of approximately 5 degrees in the grating direction and 30 degrees in the transverse direction, Fig. 1d, and an average peak efficiency of 21.5%. Due to reflections from the silicon substrate, the peak efficiency varies over the couplers, Supplemental Fig. 1. To minimize the effect of stray light and reflections, the photodiodes and area surrounding the grating couplers is covered by metal and an absorptive germanium layer. By sweeping the grating pitch, the maximum reception angle is swept from  $-10$  to  $30$  degrees (to normal) in 2 degree increments, Fig. 1e.

The fabricated device is shown in Fig. 2a. The reception patterns of 17 functional gratings were measured with a super luminescent diode (SLD) with a 50nm bandwidth, and are shown in Fig. 2b. To verify the functionality and linearity of the sensor, we imaged light impinging on a double slit and each of its constituents, Fig. 2c,d. By imaging two slits and their combination, we can test the approximation in Eq. 6, as the image formed by the combination should be the sum of the constituents. To do this, light from the SLD is first projected from a cleaved



**Figure 2.** (a) Diagram of silicon on insulator platform with germanium photodiodes (top left). Scale isometric rendering of fabricated device from angle, sides, and top. Silicon shown in purple, germanium in green, and metal connections, pads, and vias in grey. (b) Measured angular reception patterns of the 17 functional sensors. Three sensors were not functioning due to photodiode connections. (c) Representation of test setup. The chip is illuminated through a cleaved optical fiber, diffuser, and target. The copper target blocks light from some the angles it is present, forming a negative of the barcode image seen by the chip. The target and chip are mounted to a rotary stage so that the light spread from the diffuser can be averaged to effectively provide an even illumination across all visible angles. (d) Data collected (top row) and reconstruction (bottom row) of three targets: two with slits in different positions (left column) and one with both slits (right column). (Graphics made by authors).

optical fiber onto a diffuser to perturb the spacial coherent of the beam. The diffuse light then travels through a copper complementary image (negative) of a target image and impinges on the chip. The target and chip are positioned on a rotary stage so that the illumination from the diffuser can be integrated over the entire angular sensitivity range (45 degrees). The final images can then be reconstructed by inverting the measurement in Fig. 1b, our new  $\mathbf{H}$ , and multiplying by the data collected.

**Robust reconstruction algorithm.** To capture images with this sensor, we must reconstruct the unknown intensity vector  $\mathbf{x}$ , from the (noisy) linear system (8). This reconstruction problem faces two main challenges. First, because we need the discrete approximation in (6) to be accurate, the number of measurements  $m = 17$  is much smaller than the underlying dimension  $n = 120$ . Second, our algorithm must be tuned properly to adjust for the impact of the noise in the linear system.

To address both of these challenges, we employ regularized least squares<sup>16,17</sup>, which provides us with the following estimate,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \lambda \|\mathbf{\Gamma}\mathbf{x}\|^2. \quad (9)$$

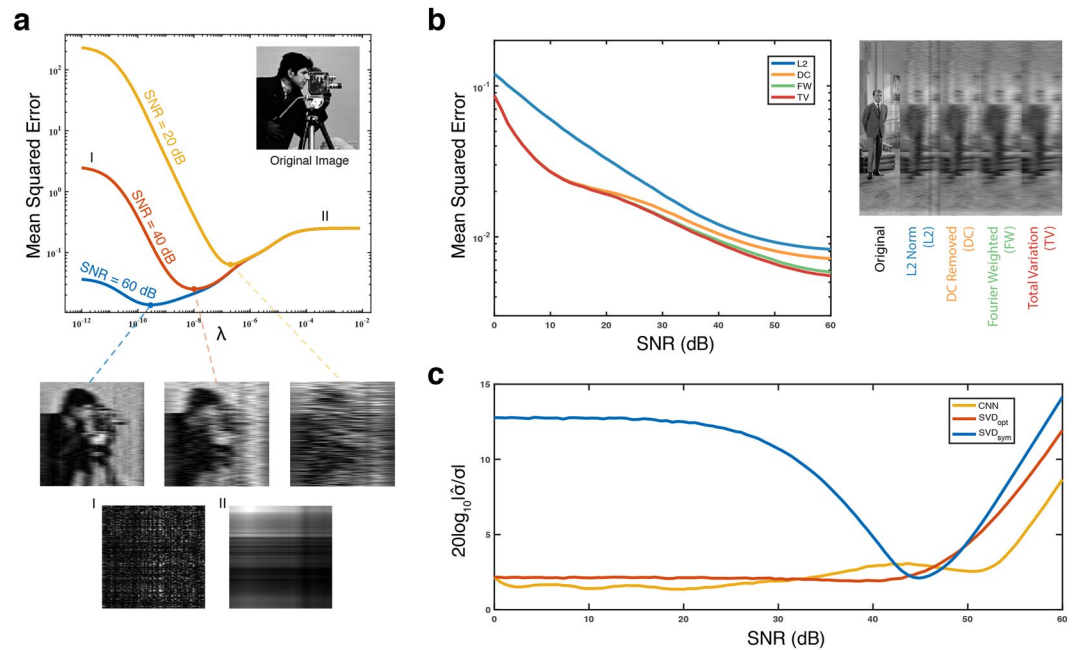
where  $\mathbf{\Gamma}$  is an  $n \times n$  matrix that is chosen based on the structure of the underlying dataset and  $\lambda$  is the regularization parameter that must be tuned properly. Note that the objective function in (9) consists of two terms. The first term in the objective function, the least-squares term, essentially captures the power of the noise vector  $\mathbf{z}$ . The second term in the optimization, the regularization term, enforces a structure dictated by  $\mathbf{\Gamma}$  on the unknown vector  $\mathbf{x}$ .

For any set of parameters, we can represent the solution to the optimization (9) in the following closed-form,

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{H} + \lambda \mathbf{\Gamma}^T \mathbf{\Gamma})^{-1} \mathbf{H}^T \mathbf{y}. \quad (10)$$

which leads to very computationally efficient solutions.

To generate an optimal reconstruction, we must choose an optimal regularization strength  $\lambda$  and regularization matrix  $\mathbf{\Gamma}$ . The choice of these parameters is critical, as can be seen in Fig. 3a.



**Figure 3.** (a) Comparison of simulated reconstruction error across  $\lambda$  for data with different SNRs. Images shown for  $\lambda_{opt}$  of each SNR as well as in the low lambda and high lambda regimes. As this is a 1D imager, an image consists of a single slice of  $1 \times 120$  pixels. The 2D images are representations of the concatenated 1D slices. (b) Comparison of reconstruction error with different  $\Gamma$  matrices at  $\lambda_{opt}$ . (c) Comparison of SNR estimators: CNN, SVD with a heuristic filter ( $SVD_{sym}$ ), and SVD with an optimized filter ( $SVD_{opt}$ ). (Public domain images made available by the University of Granada Computer Vision Group).

The regularization matrix,  $\Gamma$ , should be chosen based on the structure of the underlying data<sup>18</sup>. As images do not follow any inherent rigid structure, there is no strictly optimal  $\Gamma$ , but we can use general properties of images to choose better matrices. Here, we utilize four matrices for  $\Gamma$  which have desirable properties for image reconstruction (detailed in Methods: Gamma Matrices): The identity or L2 norm (L2), DC removed (DC), Fourier Weighted (FW), and Total Variation (TV). Figure 3b shows a comparison of the reconstruction algorithms for the above choices of the matrix  $\Gamma$ .

After choosing a favorable regularization matrix,  $\Gamma$ , we must tune  $\lambda$  to its optimal value,  $\lambda_{opt}$ . In the optimization program (9), it can be observed that the first term,  $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$ , measures the power of the noise vector, and the second term, i.e.,  $\|\mathbf{H}\mathbf{x}\|^2$ , measures the power of the signal. Hence, we expect the optimal value of the  $\lambda$  to be inversely proportional to the signal-to-noise ratio (SNR), i.e.,  $\lambda_{opt} \propto \frac{1}{SNR}$ . Extensive numerical simulation (see Supplemental Info) validates this fact. Thus to compute  $\lambda_{opt}$ , we need to evaluate the SNR.

While the noise level could be calibrated, chip-to-chip variation, temperature fluctuation, varying levels of electrical interference, and numerous additional environmental factors can affect the noise present in a measurement. As optimal reconstruction is highly sensitive to SNR, to create a more robust system, we devise two algorithms to estimate the noise variance in a given measurement. Note that the noise in the system of equations (8) is caused by two different sources: 1) The measurement noise stemming from the device. Note that this noise can have impact of both  $\mathbf{H}$  and the measurement vector  $\mathbf{y}$ . 2) The approximation error caused by discretizing the model in (6). Note that in order to have a better approximation in the discrete model, one needs to increase the signal dimension  $n$ ; however, this would have a negative impact on the reconstruction as  $\frac{m}{n}$  decreases. For our analysis purposes, we assume that the additive noise,  $\mathbf{z}$ , has independently and identically distributed Gaussian entries with mean zero and variance equal to  $\sigma^2$ . In order to find an estimate for SNR we employ two different algorithms. Our first proposed algorithm relies on the assumption that the underlying signal  $\mathbf{x}$  is smooth. This assumption is commonly used in imaging applications<sup>18</sup>. Consider the singular value decomposition of the measurement matrix  $\mathbf{H}$  as follows,

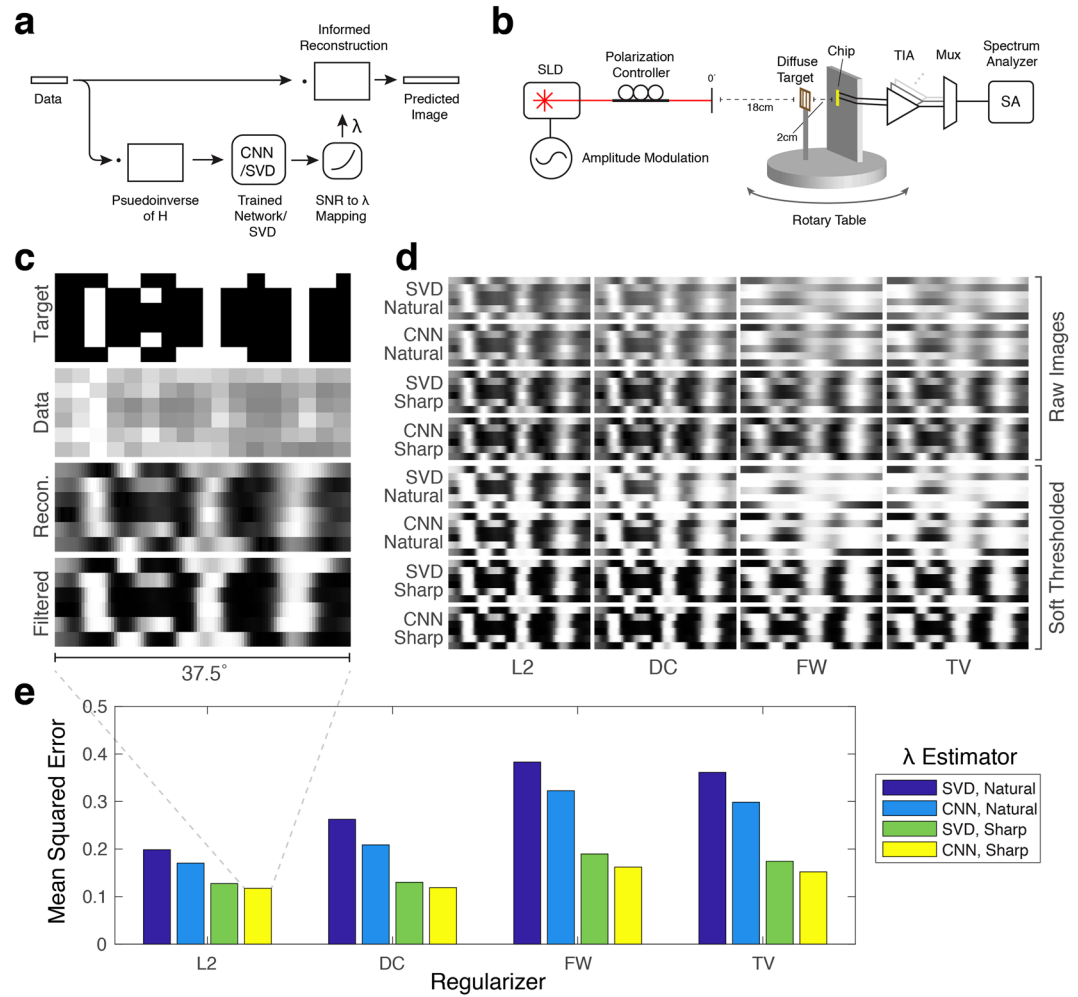
$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \quad (11)$$

where  $\mathbf{U}$  is an  $m \times m$  orthogonal matrix whose  $i^{\text{th}}$  column is represented by  $\mathbf{u}_i$ ,  $\mathbf{\Sigma} = \text{diag}(s_1, s_2, \dots, s_m)$ , and  $\mathbf{V}$  is an  $m \times n$  matrix with orthonormal columns, where  $\mathbf{v}_i$  denotes its  $i^{\text{th}}$  column. We then consider the impact of the reconstruction algorithm on the noise vector. We have,

$$\mathbf{z}^\dagger = \mathbf{H}^\dagger \mathbf{z} = \sum_{i=1}^m s_i^{-1} (\mathbf{u}_i^T \mathbf{z}) \mathbf{v}_i, \quad (12)$$

where  $\mathbf{H}^\dagger$  is the pseudo-inverse of  $\mathbf{H}$ . If we filter out the low frequency components of this reconstruction (assuming that the underlying signal is smooth), we are left with an approximation of the reconstruction contribution of noise, and  $\sigma^2$  can be approximated as,





**Figure 4.** (a) Block diagram of reconstruction algorithm. (b) Measurement setup. (c) Reconstruction example with measured CIT barcode patterns. 'Target' shows the barcodes used, displayed horizontally and concatenated vertically. 'Data' shows the data collected from each of the 17 sensors. 'Recon.' shows the image reconstructed using L2 regularization, CNN noise estimation, and an SNR- $\lambda$  mapping for sharp images. 'Filtered' shows the reconstruction after passing through a sigmoidal amplitude filter. (d) Reconstructed images using different  $\Gamma$  matrices, SNR estimators, and SNR to  $\lambda_{opt}$  mappings. 'Natural' is using the SNR to  $\lambda_{opt}$  mapping optimized for natural images, and 'Sharp' is using the SNR to  $\lambda_{opt}$  mapping optimized for random binary images. (e) Quantified reconstruction error for plots in d. Each bar represents the average of the 6 barcode patterns. (Graphics made by authors).

$$\hat{\sigma}^2 = \frac{1}{m} \sum_{i=1}^m (s_i \mathbf{v}_i^T \mathbf{z})^2. \quad (13)$$

The second proposed algorithm is to train a convolutional neural network<sup>19,20</sup>. This approach is particularly useful when we have enough samples from the dataset. Unlike the first approach, here we do not need any assumption on the structure of the underlying signal. In our implementation, we trained a 4-layer network with two convolution layers and two fully-connected layers. The details of the training is deferred to Methods. Figure 3c compares the result of the above-mentioned algorithms in estimating the SNR. As expected, the CNN approach often provides a better estimate.

To form an image, the data collected from the sensor is fed through the SNR estimator, the SNR is mapped to  $\lambda_{opt}$ , and the data is finally processed with the reconstruction (9), Fig. 4a.

**Measurement.** To test the full system, we measure 6 barcode patterns, Fig. 4c (Target), that spell the letters C, I, T when concatenated. The collected sensor data is then multiplied by a reconstruction matrix  $\mathbf{H}$  and fed through the trained CNN or SVD algorithm to estimate the noise variance. This noise variance is then used to map to an optimal lambda estimate through SNR. Finally, a new  $\mathbf{H}_{opt}$  is calculated using this optimal lambda and the image is reconstructed, Fig. 4a. Figure 4c–e show the resulting images reconstructed using each of the four  $\Gamma$  matrices, SVD and CNN algorithms, and two SNR to  $\lambda$  mappings. We achieved a minimum mean squared error of 11.7%, equivalent to a peak SNR of 9.3 dB.

## Discussion

Here we develop an imaging technique that utilizes an engineered set of grating couplers in integrated photonics to capture images without the need for a lens or active beam-steering. We demonstrate this technique with a 1-dimensional imager that occupies only  $2000 \times 200 \times 20 \mu\text{m}^3$  and can capture 20 data points a 45-degree field of view. This type of miniature image sensor could enable imaging in systems that require extremely light-weight and compact sensors, such as wearable electronics, drones, and health monitors. In addition, the demonstrated algorithms can be used in other computational imaging systems, and in many cases can easily coordinate with existing infrastructure using linear regularizers to improve robustness to noise.

Because it is implemented in an integrated photonics process, this type of sensor can be augmented by photonic processing techniques such as filtering and passive beam-forming<sup>21–23</sup>, and can be generalized to 2D and 3D imaging. By optimizing gratings for transverse as well as longitudinal angles, the same type of imaging can be applied to a 2D field of view, and by using parallax or optimization of near field reception patterns, this type of imager could also capture 3D data. For high resolution in the extra dimensions, the imager system requires many more gratings, and thus larger sensors for a similar effective aperture. To get around this problem, it may be necessary to use grating couplers optimized for multiple gratings to increase the camera's effective aperture. To reconstruct these 2D and 3D images, we can take the same algorithmic approach as in the 1D case by adding the extra dimensions in the transfer matrix  $\mathbf{H}$  (replacing it with a tensor or flattened tensor). While the reconstruction will be more computationally intensive, the longer computational time can be shortened using pre-calculated matrix inverses. In addition, photonics information processing techniques could be used in as a preprocessing step or in conjunction with the reconstruction algorithm<sup>24,25</sup>.

## Methods

**Sensor characterization.** The angular sensitivity of each grating coupler-photodiode pair was measured by recording the photodiode output current in each of 120 angles as set by a rotating table the chip was mounted to. The light source (laser or SLD) was projected from a cleaved fiber approximately 20 cm from the chip surface. To improve SNR, the laser/SLD was modulated with a continuous wave and the amplified output current was read by a spectrum analyzer.

**Imaging targets.** Targets were imaged as shown in Fig. 4b. Amplified and polarized SLD light was projected onto a ground glass and copper diffusive target with an optical power density of  $0.51 \text{ mW/cm}^2$ , illuminating the chip. The polarization of the projected light was aligned with the grating couplers to allow for maximum reception efficiency. Output power was measured for each of the 17 functional grating coupler-photodiode pairs in each of 40 angles spanning 50 degrees. These measurements were then averaged, effectively modeling a uniform illumination over the 45 degree field of view.

**Simulation data.** To test the reconstruction algorithms, 18432 training and 6144 test images were used, formed by slicing  $48 \times 512 \times 512$  pixel images (graciously collected and made public by the Computer Vision Group at the University of Granada) into  $1 \times 120$  pixel 1D images. To simulate data from the sensor, these images were multiplied by the measured  $\mathbf{H}$  matrix and i.i.d. Gaussian noise was added to achieve the desired SNR.

**Gamma matrices.** Here, we utilize the following matrices for  $\Gamma$ :

1.  $\Gamma = \mathbf{I}_n$ : This corresponds to least-squares with  $\ell_2$  regularization. It has been observed that this regularization is useful as for stabilizing the solution of the least-squares.
2.  $\Gamma = \mathbf{F}^T \mathbf{D} \mathbf{F}$ : Here  $\mathbf{F}$  is the  $n \times n$  discrete Fourier transform and  $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_n)$  is a diagonal matrix. This choice of regularization matrix is particularly useful when we have some prior information about the frequency structure of the underlying signal. In our experiments we have exploited the following two choices of the matrix  $\mathbf{D}$ :
  - $d_i = \begin{cases} 0, & \text{for } i = 1 \\ 1, & \text{for } i \neq 1 \end{cases}$ : This choice of the matrix  $\mathbf{D}$  simply removes the DC-term of the signal in the regularization.
  - $d_i = \begin{cases} 0, & \text{for } i = 1 \\ (i + 60/59)^2, & \text{for } 1 < i \leq 60 \\ ((121 - i) + 60/59)^2, & \text{for } 60 < i \leq 120 \end{cases}$ : This choice of the matrix  $\mathbf{D}$  imposes different weight on different frequency components of the underlying signal such that the high-frequency components (indices close to  $\frac{n}{2}$ ) have higher weights.
3.  $\Gamma = \Gamma_{TV}$ : This regularization is commonly used to enforce smoothness in imaging applications. The matrix for total variation regularization is defined as follows:

$$\Gamma_{TV}[i, j] = \begin{cases} 1, & \text{if } i = j \\ -1, & \text{if } i - 1 = j, \text{ for } 1 \leq i, j \leq n. \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

**Finding the optimal  $\lambda$ .** The optimal lambda for a given SNR and  $\Gamma$  was determined by numerically minimizing the mean squared error of a the 6144 test images using the Broyden-Fletcher-Goldfarb-Shanno<sup>26</sup> algorithm.

**SVD algorithm.** In order to achieve the optimal filter weights for the SVD algorithm (described in Section 3.2), we numerically minimized the loss function  $mean\left(\left|log_{10}\left(\frac{\hat{\sigma}}{\sigma}\right)\right|\right)$ , where  $\sigma$  denotes the power of the noise, and  $\hat{\sigma}$  is its estimate. We trained the Broyden-Fletcher-Goldfarb-Shanno algorithm across the 18432 training images with logarithmically distributed  $\sigma$  such that the SNR varies from 0 dB to 60 dB. After doing the training algorithm, we found an optimal filter of size 5 and weights equal to  $[-3.365, 6.881, -0.506, -3.417, 1.402]$ . As depicted in Fig. 3c, exploiting the optimal filter for SVD improves the performance significantly compare to the symmetric filter  $\left(-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right)$ .

**CNN algorithm training.** We trained a Convolutional neural network with the following architecture:

- Layer 1: 1D Convolution; 30 filters of size 51, activation: sigmoid, padding: same size. 1D Max Pooling; Pool size = 4
- Layer 2: 1D Convolution; 5 filters of size 15, activation: sigmoid, padding: same size
- Layer 3: Fully Connected layer; 25 neurons, activation: sigmoid
- Layer 4: Fully Connected layer; 15 neurons, activation: sigmoid. Fully Connected neuron; 1 neurons, activation: exponential

The CNN was trained using the Adam optimizer to minimize  $mean\left(\left|log_{10}\left(\frac{\hat{\sigma}}{\sigma}\right)\right|\right)$  with logarithmically distributed  $\sigma$  across the 18432 training images.

Received: 16 November 2019; Accepted: 3 January 2020;

Published online: 03 February 2020

## References

1. Jafferis, N. T., Helbling, E. F., Karpelson, M. & Wood, R. J. Untethered flight of an insect-sized flapping-wing microscale aerial vehicle. *Nature* **570**, 491 (2019).
2. Win, S. K. H., Win, L. S. T., Sufiyan, D., Soh, G. S. & Foong, S. Dynamics and control of a collaborative and separating descent of samara autorotating wings. *IEEE Robotics and Automation Letters* **4**, 3067–3074 (2019).
3. Tseng, D. *et al.* Lensfree microscopy on a cellphone. *Lab on a Chip* **10**, 1787–1792 (2010).
4. Shechtman, Y. *et al.* Phase retrieval with application to optical imaging: a contemporary overview. *IEEE signal processing magazine* **32**, 87–109 (2015).
5. Sun, B. *et al.* 3d computational imaging with single-pixel detectors. *Science* **340**, 844–847 (2013).
6. Asif, M. S., Ayremlou, A., Sankaranarayanan, A., Veeraraghavan, A. & Baraniuk, R. G. Flatcam: Thin, lensless cameras using coded aperture and computation. *IEEE Transactions on Computational Imaging* **3**, 384–397 (2016).
7. Mait, J. N., Euliss, G. W. & Athale, R. A. Computational imaging. *Advances in Optics and Photonics* **10**, 409–483 (2018).
8. El Gamal, A. Trends in cmos image sensor technology and design. In *booktitleDigest. International Electron Devices Meeting*, 805–808 (organization IEEE, 2002).
9. Robinson, M. D. & Stork, D. G. Joint digital-optical design of superresolution multiframe imaging systems. *Applied optics* **47**, B11–B20 (2008).
10. Jalali, B. & Fathpour, S. Silicon photonics. *Journal of lightwave technology* **24**, 4600–4615 (2006).
11. Jiang, J. & Fan, J. A. Global optimization of dielectric metasurfaces using a physics-driven neural network. *Nano Lett.* **19**, 5366–5372 (2019).
12. Sapra, N. V. *et al.* Inverse design and demonstration of broadband grating couplers. *IEEE Journal of Selected Topics in Quantum Electronics* **25**, 1–7 (2019).
13. Brückner, A. *et al.* Thin wafer-level camera lenses inspired by insect compound eyes. *Optics Express* **18**, 24379–24394 (2010).
14. Song, Y. M. *et al.* Digital cameras with designs inspired by the arthropod eye. *Nature* **497**, 95 (2013).
15. Gill, P. R., Lee, C., Lee, D.-G., Wang, A. & Molnar, A. A microscale camera using direct fourier-domain scene capture. *Optics letters* **36**, 2949–2951 (2011).
16. Tikhonov, A. N. Solution of incorrectly formulated problems and the regularization method. *Dokl. Akad. Nauk.* **151** (1963).
17. Neumaier, A. Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM review* **40**, 636–666 (1998).
18. Benning, M. & Burger, M. Modern regularization methods for inverse problems. *Acta Numerica* **27**, 1–111 (2018).
19. Lawrence, S., Giles, C. L., Tsoi, A. C. & Back, A. D. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks* **8**, 98–113 (1997).
20. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097–1105 (2012).
21. Horst, F. *et al.* Cascaded mach-zehnder wavelength filters in silicon photonics for low loss and flat pass-band wdm (de-) multiplexing. *Optics express* **21**, 11652–11658 (2013).
22. Fatemi, R., Abiri, B. & Hajimiri, A. A one-dimensional heterodyne lens-free OPA camera. In *booktitle CLEO: Science and Innovations*, STu3G-3 (organization Optical Society of America, 2016).
23. Fatemi, R., Abiri, B. & Hajimiri, A. An  $8 \times 8$  heterodyne lens-less OPA camera. In *booktitle CLEO: QELS\_Fundamental Science*, JW2A-9 (organization Optical Society of America, 2017).
24. Vandoorne, K. *et al.* Experimental demonstration of reservoir computing on a silicon photonics chip. *Nature communications* **5**, 3541 (2014).
25. Shen, Y. *et al.* Deep learning with coherent nanophotonic circuits. *Nature Photonics* **11**, 441 (2017).
26. Shanno, D. F. Conditioning of quasi-newton methods for function minimization. *Mathematics of computation* **24**, 647–656 (1970).

## Author contributions

A.W., P.K. and A.H. conceived of and designed the imaging device. A.W., P.K., F.S., B.H. and A.H. designed the computational reconstruction algorithms. All authors contributed to writing the manuscript.



## Competing interests

The authors have filed a patent application based on the research in this paper.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-58027-1>.

**Correspondence** and requests for materials should be addressed to A.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020